## REMARKS

Reconsideration of the present application is respectfully requested. Claims 2-6 and 75-87 are pending in the application. Claims 2-4 have been amended. Claims 75-87 have been added. Claim 1 has been cancelled, claims 77 and 81 represent parts of original claim 1. Claim 77 represents former claim 1 part (c), claim 81 represents former claim 1 part (b). No new matter has been added in this amendment.

## ELECTION

The election of Group I (claims 1-6) with traverse has been made final. Claims 7-74 are non-elected and withdrawn from consideration.

**Double Patenting Rejection**

Claims 2-6 are rejected under the judicially created doctrine of obviousness-type double patenting as being unpatentable over claims 1-10 of U.S. Patent No. 6,232,527.

Applicant respectfully disagrees. The current application is a divisional application filed in response to the Restriction Requirement of U.S. Application Serial No. 09/426,557, now U.S. Patent No. 6,232,527. Claims 2-6 are not obvious in light of claims 1-10 of U.S. Patent No. 6,232,527. Claim 1 has been cancelled. Claim 77 represents the former claim 1(c), and claims an isolated polypeptide comprising an amino acid sequence having at least 80% sequence identity over the entire length of SEQ ID NO: 2, wherein the amino acid sequence encodes a polypeptide having flap endonuclease activity. Claims 3 and 4 have been amended and now depend from claim 77 (equivalent to original claim 1 (c), which was not included in this rejection), therefore claims 3-6 are removed from the double patenting rejection. Claim 2 is directed to a recombinant expression cassette which expresses the polypeptide of claim 77 (former claim 1(c)). Therefore, the claim is defined by the polypeptide which was restricted out of U.S. Application Serial No. 09/426,557 as a separate

invention. The claim is consonant with the polypeptide claim restricted, therefore it is respectfully requested that the double patenting rejection be withdrawn.

**Rejection under 35 U.S.C §101 – Utility**

Claims 1-6 are rejected under 35 U.S.C §101 because the claimed invention is not supported by either a specific and credible asserted utility or a well-established utility.

Applicants respectfully disagree, page 2, lines 3-31 of the specification clearly details the well-established utility of Rad2/FEN-1 as a structure-specific DNA endonuclease involved in DNA excision repair. Rad2/FEN-1 is involved in DNA replication, nucleotide excision repair, double strand break repair, end joining, and strand exchange during homologous recombination. The present invention proposes to use the well-established utility of Rad2/FEN-1 in order to modulate repair activity, recombination, gene targeting, or provide a male sterile phenotype in plants (see page 2, lines 22-31) therefore establishing specific, substantial and credible utility for the present invention.

The Examiner asserts for claim 1 (a) "It is unclear as to how this sequence would be useful, lacking its functional activity". The Examiner further states that claim 1 (b)-(d) lack a recitation of function.

Claim 1 has been cancelled. New claims 77 and 81 represent parts of original claim 1, and each recite "wherein the polypeptide has flap endonuclease activity". Support for this amendment is found on page 2, lines 3-17 of the specification. Claim 77 recites a polypeptide having at least 80% sequence identity over the entire length of SEQ ID NO: 4. New claims 78-80, which depend from claim 77, recite 85, 90, and 95% sequence identity respectively. Support for this amendment can be found on page 48, lines 23-29 of the specification. Therefore, the claims now recite the functional activity.

Besides the sequence comparison shown in Example 4, pages 72-73 which points out conserved regions and amino acids targeted by site-directed

mutagenesis, multiple sequence alignments are presented in Appendix A, and further sequence analysis is presented in Appendix B. These analyses show the overall homology and conserved domains shared by the claimed sequences and other Rad2/FEN-1 sequences. Further, these analyses could be used to predict protein regions likely to be tolerant of amino acid substitutions and also indicate likely acceptable substitutions.

Appendix A presents two multiple sequence alignments performed using the PileUp program in the GCG sequence analysis suite. The first alignment shows SEQ ID NO: 2 aligned with Rad2/FEN-1 homologues from yeast (*S. pombe*), mouse and human. The second alignment contains SEQ ID NO: 4 aligned with the same sequences, SEQ ID NO: 4 was also shown in Example 4 of the specification. The alignments indicate homology to Rad2/FEN-1 sequences across the entire length of SEQ ID NO: 2 and 4, and that SEQ ID NO: 2 and 4 show a high level of conservation with the N, I, and C regions of the Rad 2 family, as seen by previous workers (for example, see Ref A4 of IDS submitted 1/5/00 for Serial No. 09/426,557, now U.S. Patent 6,232,527).

Appendix B presents sequence analyses performed by the BioScout suite of programs and includes the summary report, the 3D structure alignment to a FEN-1 crystal structure PDB 1B43 (Hosfield *et al.*, 1998 Cell 95:135-146, Ref. A7 of IDS submitted 1/5/00 for Serial No. 09/426,557, now U.S. Patent 6,232,527), Pfam analyses results which indicate the presence of three Rad2/FEN-1 Pfam domains (e values = 1.2e –44, 1.3 e –36, and 6.9e -5), the alignments of SEQ ID NO: 4 with the Pfam domains, descriptions of the Pfam domains, and the top five BLOCKS hits and scores. Therefore, Applicant has shown the claimed sequences comprise extensive homology and conserved functional domains with known Rad2/FEN-1 sequences.

Support for polypeptides having a given percent sequence identity is taught in the specification as originally filed, for example, in SEQ ID NOS: 1-8; guidance on codon preferences (page 7, lines 22-31, and page 62, line 27 – page 63, line 16); codon degeneracy (page 6, lines 1-19); sequence comparisons (page 17, line 21 –

page 23, line 9, and page 64, line 19 – page 67, line 23); conservative substitutions and variants (page 6, line 20 - page 7, line 11, page 48, lines 11-29); polypeptide expression (page 41, line 1 – page 46, line 23, page 49, line 20 – page 52, line 31, and page 55, lines 15-28, page 62, line 27 – page 63, line 16), isolation (page 55, line 30 – page 56, line 18) as well as the Examples.

Assays for Rad2/FEN-1 activity were known in the art at the time of filing. For example, measuring endonuclease activity on ss circular M13 DNA (Habraken *et al.*, 1993 *Nature* 366:365-368 Ref. A2 in IDS submitted January 5 2000 for Serial No. 09/426,557, now U.S. Patent 6,232,527), structure-specific (flap) endonuclease activity (Harrington and Lieber, 1994 EMBO J 13:1235-1246 Ref. A3, Kimura *et al.* 1997 *NAR* 25(24):4970-4976 Ref. A6, Alleva and Doetsch 1998 *NAR* 26(26):3645-3650 Ref. A9, all in IDS submitted January 5, 2000 for Serial No. 09/426,557, now U.S. Patent 6,232,527), 5'-3' exonuclease activity (*e.g.*, Ref. A3, Ref. A6., and Ref A9). Also available were complementation tests, gene knockouts, UV sensitivity, cell cycle, and replication assays. Therefore, one of skill in the art could readily identify macromolecules having Rad2-FEN-1 activity.

Applicants believe that the claimed invention has a well-established utility for which they have proposed specific, substantial and credible uses. Therefore, it is respectfully requested that the rejection of claims 1-6 under 35 U.S.C §101 should be withdrawn and not applied to new claims 75-87.

**Rejection under 35 U.S.C §112, 1st paragraph – Enablement**

Claims 1-6 are rejected under 35 U.S.C §112, first paragraph since the claimed invention has been rejected under 35 U.S.C §101 for lacking utility, one skilled in the art would not know how to use the claimed invention.

As the Applicants believe the response to the utility rejection under 35 U.S.C. §101 has overcome that rejection, therefore the concomitant rejection of claims 1-6 under 35 U.S.C. §112, first paragraph based on a lack of utility should be withdrawn, and should not be applied to new claims 75-87.

Claims 1-6 are rejected under 35 U.S.C §112, first paragraph as containing subject matter which was not described in the specification in such a way to enable one of skill in the art to make and/or use the invention.

The Examiner asserts the specification only teaches that a Rad2 gene of yeast was known. Examiner also states that there is no evidence provided that SEQ ID NO: 2 has similar structural features to the yeast Rad2 gene, or evidence of functional activity of the claimed polypeptides, or guidance on how to evaluate SEQ ID NO: 2 for a specific activity. The Examiner asserts that undue experimentation would be required to make and/or use the invention.

Applicants respectfully disagree. As it is stated in the specification on page 2, lines 3-22 not only was the yeast gene known, but other members of the gene family were identified from human (XP-G, FEN-1), budding yeast, mouse and frog (*Xenopus*). Also, FEN-1 activity had been purified from cauliflower by Kimura *et al.* (Ref. A6). Applicants disclosed several Rad2/FEN-1 polypeptide sequences in SEQ ID NOS: 2, 4, 6, and 8. Example 4 of the specification shows an alignment of SEQ ID NO:4 and human Rad2/FEN-1 (SEQ ID NO: 10) which indicates an overall sequence homology and further points out conserved XPG motifs, and residues involved in catalysis and substrate binding. The overall homology and conserved motifs and amino acids are evidence of functional activity. Further evidence has been presented in Appendices A and B as discussed above. Assays for Rad2/FEN-1 activity were known in the art, as discussed above. Further, these references are cited on pages 2-3 of the specification and presented in the IDS submitted 1/5/00 for the parent application Serial No. 09/426,557, now U.S. Patent 6,232,527. As is stated in MPEP 2164.01 "A patent need not teach, and preferably omits, what is well known in the art. *In re Buchner*, 929 F.2d 660, 661, 18 USPQ2d 1331, 1332 (Fed. Cir. 1991); *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 USPQ 81, 94 (Fed. Cir. 1986), *cert. denied*, 480 U.S. 947 (1987); and *Lindemann Maschinenfabrik GMBH v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1463, 221 USPQ 481, 489 (Fed. Cir. 1984)."

The Examiner asserts that sequence homology is not sufficient to predict function, citing Doerks *et al.* (*TIG* 14(6):248-250 1998), Smith *et al.* (*Nature Biotech.* 15:1222-1223 1997), Brenner (*TIG* 15(4):132-133 1999), and Bork (*TIG* 12(10):425-327 1996).

Applicant respectfully disagrees. These references all point to the potential accumulation and propagation of errors from automatic computer annotation of high throughput sequencing. This kind of sequencing and annotation was frequently a one-pass sequencing and annotation by top local alignment hit. These do not state that analyses of sequence homology are not predictive of function, these point to the problem of automatic annotation based on one top hit in a database. Doerks *et al.* actually try to predict the function of uncharacterized protein families (UPFs) through sequence analyses alone (page 248, column 1-3). Smith points out that the problems do not lie in the homology search algorithms, but in minor database annotation inconsistencies (page 1222, col. 1, 3rd paragraph – col. 2, 1st paragraph). Brenner's statement that ".. then most homologs must have different molecular and cellular functions" is prefaced by "**if** there are only about a 1000 major superfamilies in nature" (emphasis added). Brenner examined the annotations by three groups for the complete genome of *M. genitalium*, and estimated the minimum error rate of annotations to be 8%, therefore the correct annotation rate would be approximately 92%. Bork states "we wish to point out that sequence databases are the most useful tool in sequence analysis and the question should be how can one further improve their value" (page 472, col. 1, 1st full paragraph). Taken as a whole, these references warn of the potential error of relying on one automated annotation to identify a gene homologue, but do not state that homology is NOT indicative of function.

Examiner cites Van de Loo et al. (PNAS 92:6743-6747 1995) as an example that gene activity cannot be determined merely by similarity of sequences. Examiner states that the reference teaches that a change in only four amino acids will convert a desaturase gene to a hydroxylase gene.

Applicant respectfully disagrees. It was the similarity of the activities, cofactors and substrates that suggested to van de Loo *et al.* that there might be sequence similarity as well (page 6743, col. 1-2). The hydroxylase could not be purified by biochemical means, therefore van de Loo *et al.* used differential screening and sequencing to identify seed-specific clones that were similar to desaturase. A hydroxylase cDNA clone was identified and confirmed based on its similarity to a desaturase. Applicant found no teaching of amino acid changes to convert a desaturase to a hydroxylase in this reference.

Applicants submit that no more than routine experimentation is required. This may be accomplished by the examples and methods within the present application and within the technical, scientific, skill in the art. The disclosure of SEQ ID NOS: 1-8, the guidance on conserved motifs and amino acids (*e.g.* Example 4), sequence comparisons, and modifications discussed above, and the ready availability of routine Rad2/FEN-1 assays enabled one of skill in the art to make and use the claimed invention. Therefore it is respectfully requested that the rejection of claims 1-6 under 35 U.S.C §112, first paragraph be withdrawn, and not applied to new claims 75-87.

**Rejection under 35 U.S.C §112, 1st paragraph – Written Description**

Claims 1-6 are rejected under 35 U.S.C §112, first paragraph as containing subject matter which was not described in the specification in such a way to reasonably convey to one of skill in the art that the inventor(s) had possession of the claimed invention.

The Examiner asserts that the Applicant does not identify structural features for SEQ ID NO: 2. The Examiner also asserts 80% sequence identity allows for a 20% lack of sequence identity in a region essential for protein activity, and that the claims do not specify either a particular structure or a particular function.

Applicants respectfully disagree. Applicants disclosed four Rad2/FEN-1 sequences, one of which was aligned with a human sequence in Example 4 which

points out conserved functional motifs and amino acids. Claim 1 has been cancelled. New claim 77 claims a polypeptide having at least 80% identity to SEQ ID NO: 2, wherein the polypeptide has flap endonuclease activity. New claim 81 claims the polypeptide of SEQ ID NO: 2 which has flap endonuclease activity. The claims now recite both structural (percent identity) and functional limitations (flap endonuclease activity). Support for these claims is found on page 2, lines 3-17 and page 48, lines 23-29. Applicants also provided support for sequence comparisons, homology searches, and sequence modifications and other guidance as noted in the remarks regarding the rejection under 35 U.S.C 101. Applicants submit further evidence in Appendices A and B which confirm the identity of the claimed Rad2/FEN-1 sequences.

Several different physical, structural, and function characteristics of the polypeptides of the claimed invention are sufficiently described to reasonably convey to one of skill in the art that Applicants had possession of the invention. Therefore, it is respectfully requested that the rejection of claims 1-6 under 35 U.S.C §112, first paragraph be withdrawn and not applied to new claims 75-87.

**Rejection under 35 U.S.C §112, 2nd paragraph**

Claims 1-6 are rejected under 35 U.S.C §112, second paragraph as being indefinite.

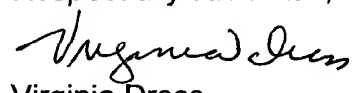
The Examiner asserts that the following terms render claim 1 indefinite: "a member", "selectively hybridize", and "stringent hybridization conditions".

Claim 1 has been cancelled, thereby obviating the rejection. The rejection of claims 1-6 under 35 U.S.C §112, second paragraph should be withdrawn and not applied to new claims 75-87.

## CONCLUSION

In light of the foregoing remarks and amendments, it is believed the claims are in condition for allowance. Entry of the amendment and withdrawal of the outstanding rejections and allowance of all of the remaining claims is respectfully requested.

Respectfully submitted,



Virginia Dress
Agent for Applicant(s)
Registration No. 48,243

PIONEER HI-BRED INTERNATIONAL, INC.
Corporate Intellectual Property
7100 N.W. 62nd Avenue
P.O. Box 1000
Johnston, Iowa 50131-1000
Phone: (515) 270-4192
Facsimile: (515) 334-6883

# APPENDIX A

!!AA_MULTIPLE_ALIGNMENT 1.0
PileUp of: @/tmp/45548148.list

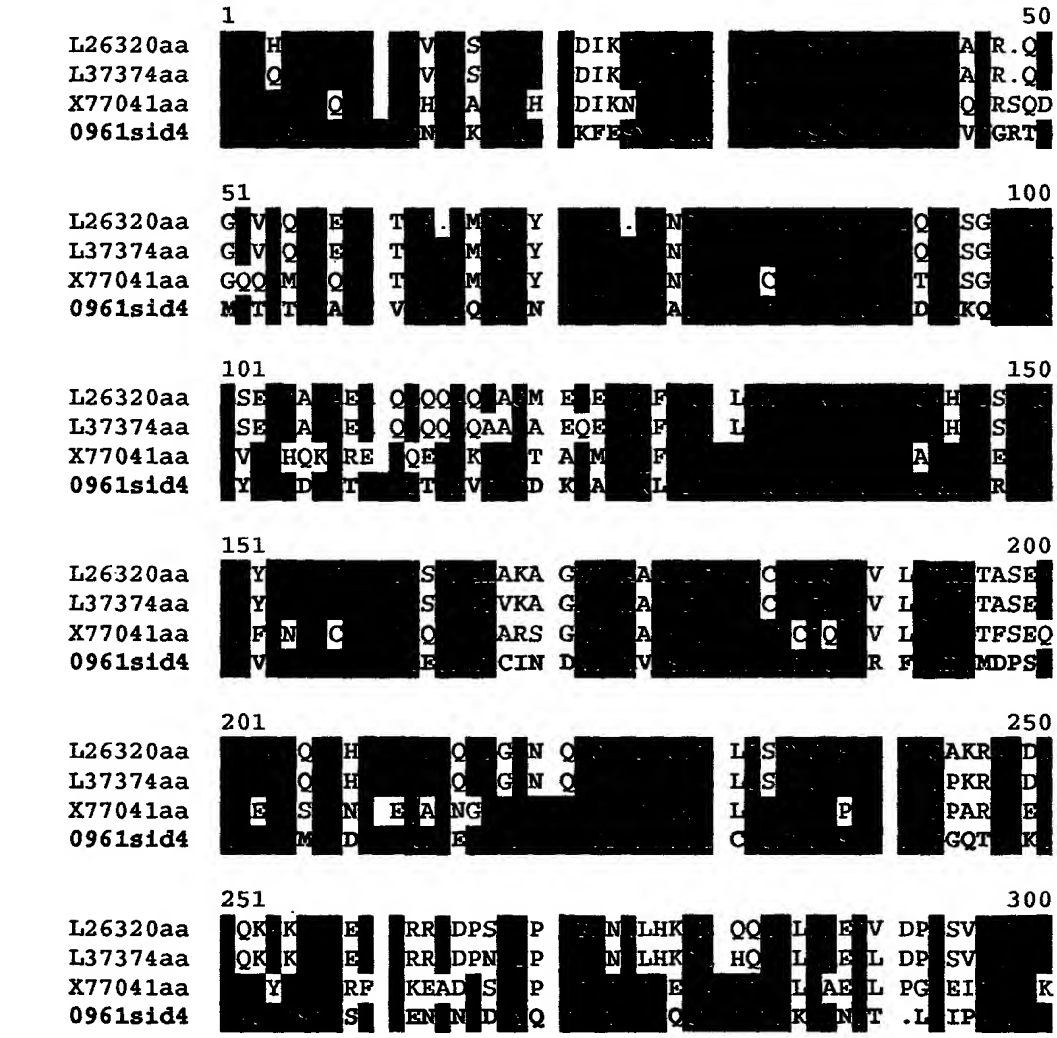
Symbol comparison table: genrundata:blosum62.cmp CompCheck: 1102

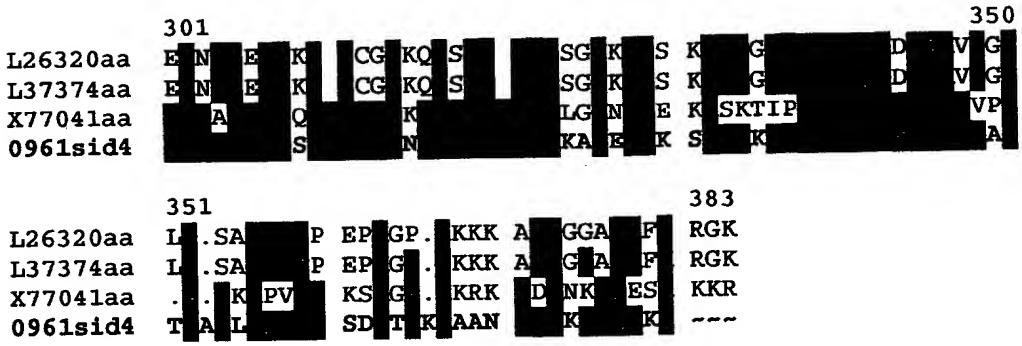
GapWeight: 8 GapLengthWeight: 2

0961L26320aa_pileup_45548.txt MSF: 383 Type: P March 5, 2003 16:16 Check: 3672 ..

| | |
|---|---|
| L26320aa | Mouse FEN-1 (Rad2) protein encoded by GenBank L26320 |
| L37374aa | Human FEN-1 (Rad2) protein encoded by GenBank L37374 |
| X77041aa | S. pombe Rad2 protein encoded by GenBank X77041 |
| **0961sid4** | **0961D SEQ ID NO: 4** |

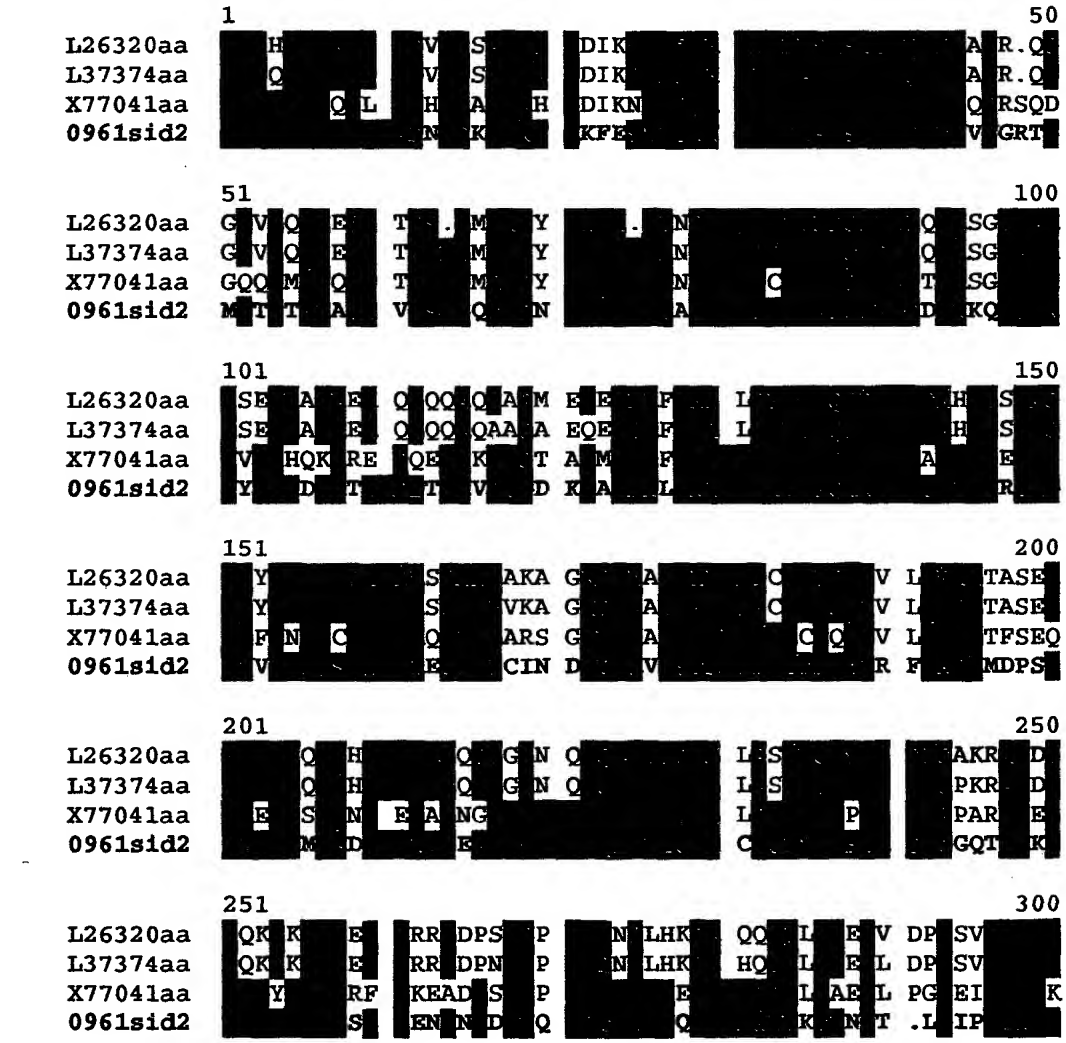
[redacted]; [redacted]

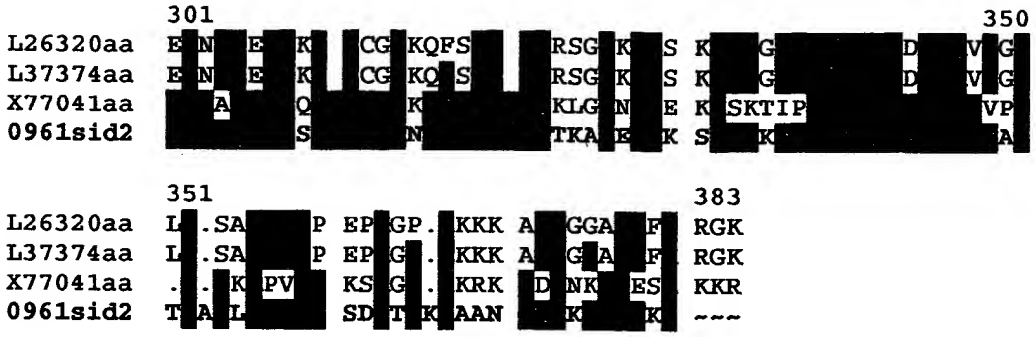




```
!!AA_MULTIPLE_ALIGNMENT 1.0
PileUp of: @/tmp/45742915.list

 Symbol comparison table: genrundata:blosum62.cmp  CompCheck: 1102

                  GapWeight: 8  GapLengthWeight: 2

 0961L26320aa_pileup_45742.txt  MSF: 383  Type: P  March 6, 2003 13:14  Check:
3748 ..

 L26320aa     Mouse FEN-1 (Rad2) protein encoded by GenBank L26320
 L37374aa     Human FEN-1 (Rad2) protein encoded by GenBank L37374
 X77041aa     S. pombe Rad2 protein encoded by GenBank X77041
 0961sid2     0961D SEQ ID NO: 2
```

[redacted]; [redacted]





U.S. Serial No. 09/805,311
Group Art Unit: 1638

# APPENDIX B

bioSCOUT®

Projects | Analyses | Admin | Alert | SRS | Tutorial | Help

Analysis Browser: Level Up

**Report for** 0961sid2.Rad2 (Protein) Update

**Description** 0961sid2_Rad2 Edit

**Function**

# DNA repair protein rad2.

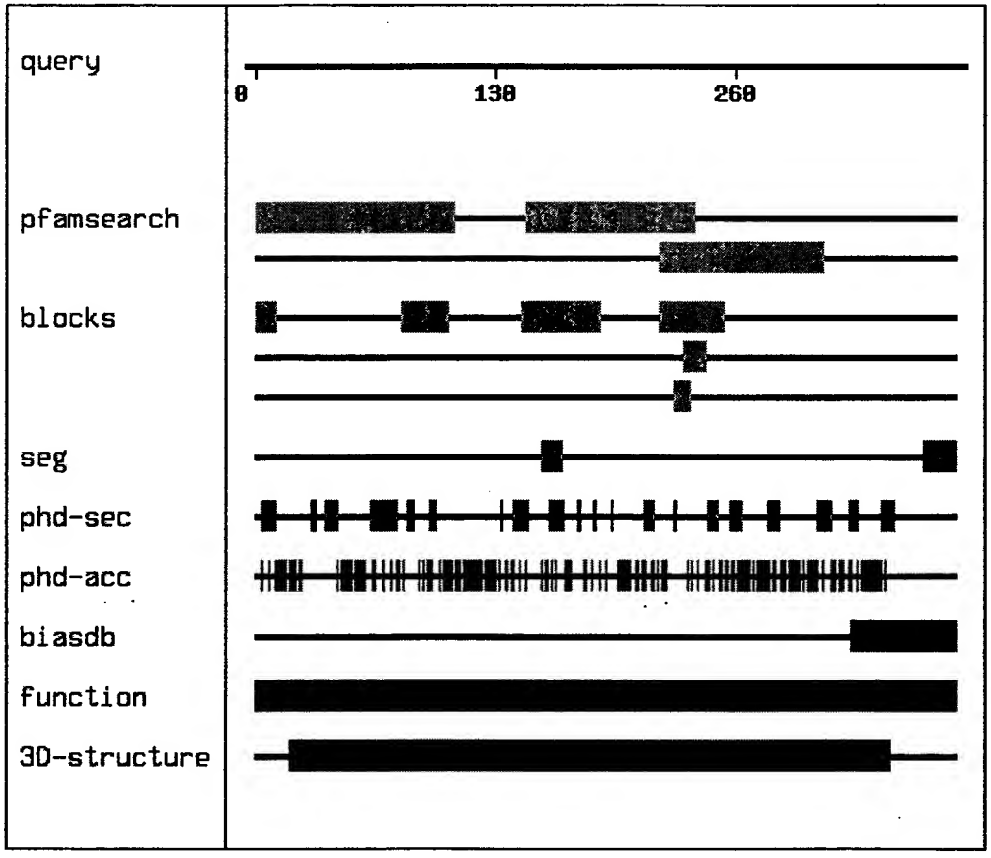
**Direct assignment of functionality by homology to** swissnew|P39750|RAD2_SCHPO

in region **1** to **379** for overall length of **380** (100% of query, 99% of hit, see the alignment ).

**Functional class** Replication

**Extracted keywords** Hydrolase, Nuclease, DNA repair, Endonuclease

**Features Summary**



| | | | |
|---|---|---|---|
| **Homologies** | | | All BLAST hits |
| | **Protein** | **125 clear** homologs | All protein BLAST hits |
| | **ESTs** | **203** homologs | All EST BLAST hits |
| | **Patents** | **199** homologs | All patent hits |
| **General** | **Gene name** | | |
| | **Molecular weight** | 42.42 kD | |
| | **Sequence length** | 379 | |
| | **Isoelectric point** | 8.52 | |
| | **Predicted cellular localisation (PHD and PreLoc)** | cytoplasmic (86.3 %) | |
| | **Identical sequence segments in:** | embl\|AI881599\|AI881599<br>genbank\|AI881599\|AI881599<br>embl\|BE639422\|BE639422<br>genbank\|BE639422\|BE639422<br>embl\|BM501417\|BM501417<br>genbank\|BM501417\|BM501417 | |
| **3D Structure** | **3D structure inferred by clear homology from residues 19 to 344 in 1B43-A** | | |
| | **View** | alignment | |
| | **pdb\|1B43\|1B43-A** | structure | |
| **Expression** | **Expression of this gene is reported for** | | |
| | **Organ Category** | other species | |
| | **Development Stage** | just after the transition from vegetative to inflorescence | |
| | **Tissue Classification** | normal tissue | |
| **Phylogeny** | **Distribution** | 45 species extracted from 263 homologous sequences. | Species |
| | **Taxa** | Archaebacteria, Chordata, Eukaryotae, Fungi, Planta, Viridae | |
| | **Model organisms** | *Arabidopsis thaliana, Caenorhabditis elegans, Drosophila melanogaster, Homo sapiens, Mus musculus, Saccharomyces cerevisiae* | |

| | | |
|---|---|---|
| **Features** | **Low complexity region** | from **155** to **166,** from **362** to **379** detected by [seg] |
| | **K-rich region** | from **323** to **379** detected by [biasdb] |
| | **No significant hits detected by** | [Coils] [Phd-tm] |
| **Patterns** | **5'-3' exonuclease, C-terminal SAM fol region** | from residue **219** to **308.** Source: [pfamsearch] . Quality: (E=6.9e-05) |
| | **XPG I-region region** | from residue **147** to **238.** Source: [pfamsearch] . Quality: (E=1.3e-36) |
| | **XPG N-terminal domain region** | from residue **1** to **108.** Source: [pfamsearch] . Quality: (E=1.2e-44) |
| | **XPG protein. region** | from residue **219** to **254.** Source: [blocks database] |
| | | from **145** to **187.** Source: [blocks database] |
| | | from **80** to **105.** Source: [blocks database] |
| | | from **1** to **11.** Source: [blocks database] |
| | **No significant hits found in** | [prosite database] |

**Comment** No comment section. Edit

**Completed Tasks**

| Start Time | User | Comment | Output | Interactiv |
|---|---|---|---|---|
| 03.03.2003, 13:30:31 | dressvm | | bioSCOUT_default details... | |

**Permissions** Edit

**Alert Jobs** New

---

Please report problems and feedback concerning bioSCOUT through the support interface.

bioSCOUT®

Projects Analyses Admin Alert SRS Tutorial Help Features

# Alignment: 0961sid2.Rad2 - pdb|1B43|1B43-A

## BLASTP - alignment of 0961sid2.Rad2 against pdb|1B43|1B43-A

fen-1

- This hit is scoring at : 2e-63 (expectation value)
- Alignment length (overlap) : 339
- Identities : 41 %
- Scoring matrix : BLOSUM62 (used to infer consensus pattern)
- Database searched : nrdb

```
Q:      19 KEQKFESYFGRKIAVDASMSIYQFLIVVGRTGMETLTNEAGEVTSHLQGMFNRTIRLLEA
           KE :.E:.:G:KIA:DA  :IYQFL .: :.. ..L.:..G.:TSHL.G:F RTI.L:EA
H:      11 KEIELENLYGKKIAIDALNAIYQFLSTIRQKDGTPLMDSKGRITSHLSGLFYRTINLMEA

           GIKPVYVFDGKPPDMKKQELAKRYSKRDDATKDLTEAVEVGDKDAIEKLSKRTVKVTRQH
           GIKPVYVFDG:PP:.KK:EL.KR ..R::A.:. .EA:E G: :...K.::R..:V...
           GIKPVYVFDGEPPEFKKKELEKRREAREEAEEKWREALEKGEIEEARKYAQRATRVNEML

           NEDCKRLLRLMGVPVVEAPSEAEAECAALCINDKVFAVASEDMDSLTFGAPRFLRHLMDP
            ED.K:LL.LMG:P:V:APSE.EA:.A :.....V:A AS:D.DSL.FGAPR.:R:L. .
           IEDAKKLLELMGIPIVQAPSEGEAQAAYMAAKGSVYASASQDYDSLLFGAPRLVRNLTIT

           SSKKIPVMEFDV---------AKVLEELELTMDQFIDLCILCGCDY-CDSIKGIGGQTAL
           ..:K:P ... V          .:VL:EL:LT.::.I:L.IL.G.DY  ..IKGIG :.AL
           GKRKLPGKNVYVEIKPELIILEEVLKELKLTREKLIELAILVGTDYNPGGIKGIGLKKAL

           KLIRQHGSIESILENLNKD---RYQIPEDWPYQEARRLFKEPNVTLDIPELKWTAPDEEG
           :::R.            :KD   ::Q ..D .....:..F .P VT D  .L W. PDEEG
           EIVRH-----------SKDPLAKFQKQSDVDLYAIKEFFLNPPVT-DNYNLVWRDPDEEG

           LISFLVKDNGFNEDRVTXAIEXIXSAXNXSSQGRLESFF     344
           ::.FL..::.F:E:RV...:E.:..A.....Q..LES:F
           ILKFLCDEHDFSEERVKNGLERLKKAIKSGKQSTLESWF     337
```

### Legend of Alignment

: positive score

. score between -2 and 0

Please report problems and feedback concerning bioSCOUT through the support interface.

bioSCOUT®

Projects | Analyses | Admin | Alert | SRS | Tutorial | Help | Features

Summary

# Searched query 0961sid2.Rad2 against PFAM database.

| Hit | Score | Expect | Description | Q from | Q to | Method |
|---|---|---|---|---|---|---|
| ☐ pfam\|hmm\|XPG_N . alignment | 161.7 | 1.2e-44 | XPG N-terminal domain | 1 | 108 | HMMPFAM |
| ☐ pfam\|hmm\|XPG_I . alignment | 135.0 | 1.3e-36 | XPG I-region | 147 | 238 | HMMPFAM |
| ☐ pfam\|hmm\|5_3_exonuclease . alignment | 9.7 | 6.9e-05 | 5'-3' exonuclease, C-terminal SAM fol | 219 | 308 | HMMPFAM |

New Task | Rename Sequences

Please report problems and feedback concerning bioSCOUT through the support interface.

bioSCOUT®

Projects Analyses Admin Alert SRS Tutorial Help Features

Alignment: 0961sid2.Rad2 - pfam|hmm|XPG_N

## HMMPFAM - alignment of 0961sid2.Rad2 against pfam|hmm|XPG_N

XPG N-terminal domain

- This hit is scoring at : 161.7
- Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:       1 MGIKGLTKLLADNAPKAMKEQKFESYFG--RKIAVDASMSIYQFLIVVGRTGMETLTNEA
           MGIKGL..:L.. AP:A::. ..E:. G  : :A:DAS: :YQFL .V  .  ..L.NE.
H:       1 MGIkGLlpiLkpvapeairsvsiEalegYYkvLAiDasiwLyqfLkavRdqlgnnlenEe

           GEVTSHLQGMFNRTIRLLEAGIKPVYVFDGKPP-DMKKQELAKRYSKRDDA     108
           GE.TSHL.G:F:R..RLL: GIKP::VFDG .P D:K.:.L.KR ::R.:A
           GettshlmglfsRlcrLldfgIkPifVFDGgapndlKaetlqKRsarrqea     111
```

### Legend of Alignment

: positive score

. score between -2 and 0

Please report problems and feedback concerning bioSCOUT through the support interface.

# Protein families database of alignments and HMMs



XPG_N

**Figure 1: 1a76**
**5'-3' exo/endo nuclease**
Flap endonuclease-1 from methanococcus jannaschii

*Accession number:* **PF00752**

## XPG N-terminal domain

### INTERPRO description (entry IPR006085)

Xeroderma pigmentosum (XP) [MEDLINE:94212451] is a human autosomal recessive disease, characterized by a high incidence of sunlight-induced skin cancer. People's skin cells with this condition are hypersensitive to ultraviolet light, due to defects in the incision step of DNA excision repair. There are a minimum of seven genetic complementation groups involved in this pathway: XP-A to XP-G. XP-G is one of the most rare and phenotypically heterogeneous of XP, showing anything from slight to extreme dysfunction in DNA excision repair [MEDLINE:93219111], [MEDLINE:94266772]. XP-G can be corrected by a 133 Kd nuclear protein, XPGC [MEDLINE:94212451]. XPGC is an acidic protein that confers normal UV resistance in expressing cells [MEDLINE:94266772]. It is a magnesium-dependent, single-strand DNA endonuclease that makes structure-specific endonucleolytic incisions in a DNA substrate containing a duplex region and single-stranded arms [MEDLINE:94266772], [MEDLINE:94376899]. XPGC cleaves one strand of the duplex at the border with the single-stranded region [MEDLINE:94376899].

XPG belongs to a family of proteins that includes RAD2 from budding yeast and rad13 from fission yeast, which are single-stranded DNA endonucleases [MEDLINE:94376899], [MEDLINE:94067324]; mouse and human FEN-1, a structure-specific endonuclease; RAD2 from fission yeast and RAD27 from budding yeast; fission yeast exo1, a 5'-3' double-stranded DNA exonuclease that may act in a pathway that corrects mismatched base pairs; yeast DHS1, and yeast DIN7. Sequence alignment of this family of proteins reveals that similarities are largely confined to two regions. The first is located at the N-terminal extremity (N-region) and corresponds to the first 95 to 105 amino acids. The second region is internal (I-region) and found towards the C-terminus; it spans about 140 residues and contains a highly conserved core of 27 amino acids that includes a conserved pentapeptide (E-A-[DE]-A-[QS]). It is possible that the conserved acidic residues are involved in the catalytic mechanism of DNA excision repair in XPG. The amino acids linking the N- and I-regions are not conserved.

| QuickGO | |
|---|---|
| **FUNCTION :** | nuclease (GO:0004518) |
| **PROCESS :** | DNA repair (GO:0006281) |

For additional annotation, see the PROSITE document PDOC00658 [Expasy | SRS-UK | SRS-USA]

| Alignment | Domain organisation |
|---|---|
| | |

| | |
|---|---|
| ◉ Seed (10) ○ Full (96)<br>Format Coloured alignment<br>Get alignment<br>Further alignment options here<br>Help relating to Pfam alignments here | ◉ Seed (10) ○ Full (96)<br>**As a Graphic** **As a Tree**<br>Zoom 0.5 pixels/aa. ☐ Bootstrap tree<br>View Graphic NIFAS Applet<br>To find out about the NIFAS tree-viewer, click here |
| **Species Distribution** | **Phylogenetic tree** |
| NEW! View alignments & domain organisation by species<br>Tree depth : Show all levels<br>View Species Tree | ◉ Seed (10) ○ Full (96)<br>Download tree ATV Applet<br>The trees were generated using Quicktree<br>To find out more about ATV phylogenetic tree-viewer click here |

| Database References | |
|---|---|
| **PDB**<br>You can find out how to set up Rasmol here | 1a76 ; 2; 101; PDB 2 Pfam Rasmol (unix) CATH-PDBSUM SCOP-UK |
| **PROSITE** | PDOC00658 [Expasy \| SRS-UK \| SRS-USA] |
| **HOMSTRAD** | XPG_NI |
| **SYSTERS** | XPG_N |
| **PANDIT** | XPG_N |

| Pfam specific information | |
|---|---|
| Author of entry | Bateman A |
| Type definition | Family |
| Alignment method of seed | Clustalw |
| Source of seed members | Pfam-B_491 (release 2.1) |

| HMMER build information | | |
|---|---|---|
| | **Pfam_ls** [Download HMM] | **Pfam_fs** [Download HMM] |
| Gathering cutoff | 3.0 3.0; | 25.0 25.0 |
| Trusted cutoff | 3.2 3.2; | 25.1 25.7 |
| Noise cutoff | -1.0 -1.0; | 18.3 15.3 |
| Build method of HMM | hmmbuild -F HMM_ls SEED<br>hmmcalibrate --seed 0 HMM_ls | hmmbuild -f -F HMM_fs SEED<br>hmmcalibrate --seed 0 HMM_fs |

For help on making stable links to this page click here

Comments or questions on the site? Send a mail to **pfam@sanger.ac.uk**

bioSCOUT®

Projects Analyses Admin Alert SRS Tutorial Help Features

Alignment: 0961sid2.Rad2 - pfam|hmm|XPG_I

## HMMPFAM - alignment of 0961sid2.Rad2 against pfam|hmm|XPG_I

XPG I-region

- This hit is scoring at : 135.0
- Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:     147 RLMGVPVVEAPS-EAEAECAALCINDKVFAVASEDMDSLTFGAPRFLRHLMDPSSK----
           RLMG:P.: AP. EAEA:CA L  .. V .:.:ED.D L.FGAPR.LR:L. ...K
H:       1 rlmGIpyIvAPgvEAEAQcayLekkglvdgiiTeDsDvLLFGaprllrnLtlsgkksgPs

           ----KIPVMEFDVAKVLEELEL---TMDQFIDLCILCGCDYCDS     238
               K:.: E.D:..:L.EL L   :.:Q.IDL.IL.GCDY.:.
           itslkveieeidlesllreLgLgklsreqLidlaiLlGcDYteG     104
```

### Legend of Alignment

```
:  positive score

.  score between -2 and 0
```

Please report problems and feedback concerning bioSCOUT through the support interface.

# Protein families database of alignments and HMMs



XPG_I

**Figure 1: 1b43 Transferase**
Fen-1 from p. furiosus

*Accession number:* **PF00867**

## XPG I-region

INTERPRO **description (entry** IPR006086**)**

Xeroderma pigmentosum (XP) [MEDLINE:94212451] is a human autosomal recessive disease, characterized by a high incidence of sunlight-induced skin cancer. People's skin cells with this condition are hypersensitive to ultraviolet light, due to defects in the incision step of DNA excision repair. There are a minimum of seven genetic complementation groups involved in this pathway: XP-A to XP-G. XP-G is one of the most rare and phenotypically heterogeneous of XP, showing anything from slight to extreme dysfunction in DNA excision repair [MEDLINE:93219111], [MEDLINE:94266772]. XP-G can be corrected by a 133 Kd nuclear protein, XPGC [MEDLINE:94212451]. XPGC is an acidic protein that confers normal UV resistance in expressing cells [MEDLINE:94266772]. It is a magnesium-dependent, single-strand DNA endonuclease that makes structure-specific endonucleolytic incisions in a DNA substrate containing a duplex region and single-stranded arms [MEDLINE:94266772], [MEDLINE:94376899]. XPGC cleaves one strand of the duplex at the border with the single-stranded region [MEDLINE:94376899].

XPG belongs to a family of proteins that includes RAD2 from budding yeast and rad13 from fission yeast, which are single-stranded DNA endonucleases [MEDLINE:94376899], [MEDLINE:94067324]; mouse and human FEN-1, a structure-specific endonuclease; RAD2 from fission yeast and RAD27 from budding yeast; fission yeast exo1, a 5'-3' double-stranded DNA exonuclease that may act in a pathway that corrects mismatched base pairs; yeast DHS1, and yeast DIN7. Sequence alignment of this family of proteins reveals that similarities are largely confined to two regions. The first is located at the N-terminal extremity (N-region) and corresponds to the first 95 to 105 amino acids. The second region is internal (I-region) and found towards the C-terminus; it spans about 140 residues and contains a highly conserved core of 27 amino acids that includes a conserved pentapeptide (E-A-[DE]-A-[QS]). It is possible that the conserved acidic residues are involved in the catalytic mechanism of DNA excision repair in XPG. The amino acids linking the N- and I-regions are not conserved.

| QuickGO | |
|---|---|
| **FUNCTION :** | nuclease (GO:0004518) |
| **PROCESS :** | DNA repair (GO:0006281) |

For additional annotation, see the PROSITE document PDOC00658 [Expasy | SRS-UK | SRS-USA]

| Alignment | Domain organisation |
|---|---|
| | |

| | |
|---|---|
| ◉ Seed (16) ○ Full (90)<br>Format: Coloured alignment<br>Get alignment<br>Further alignment options here<br>Help relating to Pfam alignments here | ◉ Seed (16) ○ Full (90)<br>**As a Graphic** **As a Tree**<br>Zoom 0.5 pixels/aa. ☐ Bootstrap tree<br>View Graphic NIFAS Applet<br>To find out about the NIFAS tree-viewer, click here |
| **Species Distribution** | **Phylogenetic tree** |
| NEW! View alignments & domain organisation by species<br>Tree depth: Show all levels<br>View Species Tree | ◉ Seed (16) ○ Full (90)<br>Download tree ATV Applet<br>The trees were generated using Quicktree<br>To find out more about ATV phylogenetic tree-viewer click here |

| Database References | |
|---|---|
| **PDB**<br>You can find out how to set up Rasmol here | 1a76 ; 140; 228; PDB 2 Pfam Rasmol (unix) CATH-PDBSUM SCOP-UK |
| **PROSITE** | PDOC00658 [Expasy \| SRS-UK \| SRS-USA] |
| **HOMSTRAD** | XPG_NI |
| **SYSTERS** | XPG_I |
| **PANDIT** | XPG_I |

| Pfam specific information | |
|---|---|
| Author of entry | Bateman A |
| Type definition | Family |
| Alignment method of seed | Clustalw |
| Source of seed members | Pfam-B_776 (release 3.0) |

| HMMER build information | | |
|---|---|---|
| | **Pfam_ls** [Download HMM] | **Pfam_fs** [Download HMM] |
| Gathering cutoff | 9.0 9.0; | 25.0 25.0 |
| Trusted cutoff | 21.6 21.6; | 29.1 26.8 |
| Noise cutoff | -4.7 -4.7; | 24.4 13.7 |
| Build method of HMM | hmmbuild -F HMM_ls SEED<br>hmmcalibrate --seed 0 HMM_ls | hmmbuild -f -F HMM_fs SEED<br>hmmcalibrate --seed 0 HMM_fs |

For help on making stable links to this page click here

Comments or questions on the site? Send a mail to **pfam@sanger.ac.uk**

bi SCOUT®



Alignment: 0961sid2.Rad2 - pfam|hmm|5_3_exonuclease

## HMMPFAM - alignment of 0961sid2.Rad2 against pfam|hmm|5_3_exonuclease

5'-3' exonuclease, C-terminal SAM fol

- This hit is scoring at : 9.7
- Scoring matrix : BLOSUM62 (used to infer consensus pattern)

```
Q:    219 LTMDQ-FIDLCILCGcDYCDSIKGIGG---QTALKLIRQHGSIESILE-NLNKDRY----
          LT :Q .ID. .L.G D .D:I.G: G    :TA.KL::::GS:E:I.E NL:K :
H:      1 ltPeQAiiDykALvG.DsSDNIPGVpGIGeKTAakLLkeyGSlEniyeiNldklkgkklr

          ------QIPEDwpYQEARRLFKE-------PNVTLDIPEL--KWTAPDEEGLIS     308
                :       :E . L ::          :.L.:.:L  : ..PD:E LI.
          GakklnekLla..gkedAflSrkLatiktDvpleltledlkFrlkppdkeqlie     111
```

### Legend of Alignment

```
:  positive score

.  score between -2 and 0
```

Please report problems and feedback concerning bioSCOUT through the support interface.

Protein families database of alignments and HMMs

Home | Keyword Search | Protein Search | Browse Pfam | DNA Search | Taxonomy | ftp | H

# 5_3_exonuclease

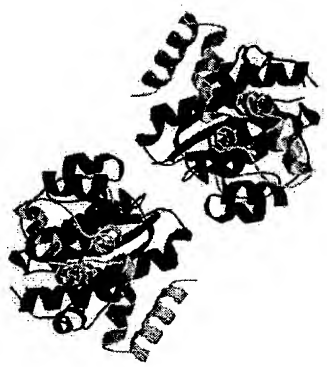


**Figure 1: 1xo1**
**Hydrolase**
T5 5'-exonuclease mutant k83a

***Accession number:*** **PF01367**

## 5'-3' exonuclease, C-terminal SAM fold

INTERPRO **description (entry** IPR002421)

The N-terminal and internal 5'3'-exonuclease domains are commonly found together, and are most often associated with 5' to 3' nuclease activities. The XPG protein signatures (PDOC00658). 5'-3' exonuclease ; GO:0008409

| QuickGO | |
|---|---|
| **FUNCTION :** | DNA binding (GO:0003677) |

| Alignment | Domain organisation |
|---|---|
| Seed (27) Full (128) | Seed (27) Full (128) |
| Format Coloured alignment | **As a Graphic** **As a Tree** |
| Get alignment | Zoom 0.5 pixels/aa. Bootstrap tree |
| Further alignment options here | View Graphic NIFAS Applet |
| Help relating to Pfam alignments here | To find out about the NIFAS tree-viewer, click here |
| **Species Distribution** | **Phylogenetic tree** |
| NEW! View alignments & domain organisation by species | Seed (27) Full (128) |
| Tree depth : Show all levels | Download tree ATV Applet |
| View Species Tree | The trees were generated using Quicktree |
| | To find out more about ATV phylogenetic tree-viewer click here |

| Database References | |
|---|---|
| **PDB** You can find out how to set up Rasmol here | 1bgx T; 174; 271; PDB 2 Pfam Rasmol (unix) CATH-PDBSUM SCOP-UK |
| **HOMSTRAD** | 5_3_exonuclease |
| | 5_3_exonuclease |

| | |
|---|---|
| **SYSTERS** | |
| **PANDIT** | 5_3_exonuclease |

| Literature References |
|---|
| 1. **A helical arch allowing single-stranded DNA to thread through T5 5'-exonuclease.** **Ceska TA, Sayers JR, Stier G, Suck D; Nature 1996;382:90-93.** |
| 2. **Structure of Taq ploymerase with DNA at the polymerase active site.** **Eom SH, Wang J, Steitz TA; Nature 1996;382:278-281.** |

| Pfam specific information | |
|---|---|
| Author of entry | Bateman A, Griffiths-Jones SR |
| Type definition | Domain |
| Alignment method of seed | Clustalw |
| Source of seed members | Pfam-B_716 (release 3.0) |

| HMMER build information | | |
|---|---|---|
| | **Pfam_ls** [Download HMM] | **Pfam_fs** [Download HMM] |
| Gathering cutoff | 7.8 7.8; | 23.0 23.0 |
| Trusted cutoff | 7.9 7.9; | 23.0 23.0 |
| Noise cutoff | 7.7 7.7; | 22.0 22.0 |
| Build method of HMM | hmmbuild -F HMM_ls SEED<br>hmmcalibrate --seed 0 HMM_ls | hmmbuild -f -F HMM_fs SEED<br>hmmcalibrate --seed 0 HMM_fs |

For help on making stable links to this page click here

Comments or questions on the site? Send a mail to **pfam@sanger.ac.uk**

Case 0961 Rad2 BioScout Analysis
BLOCKS

Top Five Hits

BLIMPS (BLocks IMProved Searcher) Version 3.2 1997/02


Probe Sequence: 0961sid2.Rad2 0961sid2_Rad2

Probe Size: 379 Amino Acids

Probe File: 0961sid2.Rad2.x

Target File (s) : /LION/data/db/native/blocks/blocks.dat

Records Searched: 4034

Scores Done: 4034

Alignments Done: 1644820

| AC# | Description | Strength |
|---|---|---|
| **BL00841D** | XPG protein. | 1585 |

| Score | RF | AA# | |
|---|---|---|---|
| 1505 | 0 | 218 | LTmdQFIDLCILcGCDYCDSIKGIGgqTALKLIRQH |

| AC# | Description | Strength |
|---|---|---|
| **BL00841C** | XPG protein. | 1629 |

| Score | RF | AA# | |
|---|---|---|---|
| 1487 | 0 | 144 | LLRLMGVPvVeAPSEAEAeCAALcinDKVfAvASEDMDsLTFG |

| AC# | Description | Strength |
|---|---|---|
| **BL00841B** | XPG protein. | 1532 |

| Score | RF | AA# | |
|---|---|---|---|
| 1476 | 0 | 79 | IKPVYVFDGKPPDmKKQELAKRySKR |

| AC# | Description | Strength |
|---|---|---|
| **BL00841A** | XPG protein. | 1260 |

| Score | RF | AA# | |
|---|---|---|---|
| 1210 | 0 | 0 | MGIKGLtKLLA |

| AC# | Description | Strength |
|---|---|---|
| **BL00167C** | 0 Tryptophan synthase alpha chain proteins. | 1274 |

| Score | RF | AA# | |
|---|---|---|---|
| 1121 | 0 | 231 | GCDYCDSIkGIgG |